

Localization and 3D Reconstruction of Urban Scenes Using GPS

Kihwan Kim, Jay Summet, Thad Starner, Daniel Ashbrook, Mrunal Kapade and Irfan Essa
College of Computing, Georgia Institute of Technology
{kihwan23,summetj,anjiro,thad,mrunal,irfan}@cc.gatech.edu

Abstract

Using off-the-shelf Global Positioning System (GPS) units, we reconstruct buildings in 3D by exploiting the reduction in signal to noise ratio (SNR) that occurs when the buildings obstruct the line-of-sight between the moving units and the orbiting satellites. We measure the size and height of skyscrapers as well as automatically constructing a density map representing the location of multiple buildings in an urban landscape.

1. Introduction

3D data sets of major buildings in urban areas are produced and sold commercially. These data sets are constructed by using expensive aerial photography, purchasing access to building permit and architectural plan databases, and performing street level laser surveying. Perhaps the most widely known use of these data sets are in GIS visualization tools such as the popular Google Earth. Here, we present an inexpensive way to produce 3D building data sets using only GPS signals.

Linked by the internet, groups of cooperating users have already begun to share GPS data for the purpose of creating free open-source maps [1]. The advantage of worn GPS enabled devices (e.g. mobile phones) is that the data collection can include relatively fine features such as walkways between buildings, footpaths in parks, and entrances to buildings. Here, we demonstrate that the addition of 3D building data-sets to these already existing efforts would only require a minimal amount of additional computation. Our technique requires many trips to collect data, and we imagine its main advantage is supporting such “crowdsourcing” or open-source data collection system. However, a motivated mobile phone service provider or vehicle fleet operator could anonymize the location information provided by their users’ GPS-enabled mobile phones or vehicle navigation systems to create a proprietary 3D map that is updated daily as its users move about a city.

In this paper we will demonstrate the feasibility of us-

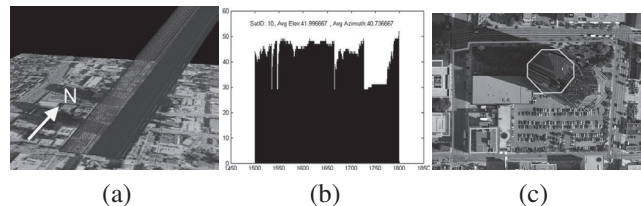


Figure 1. SNR reduction due to the obstruction of a building (Green indicates low SNR)(a)(b), Testing site(c).

ing GPS signals to determine the location, size, and height of buildings. Our approach takes advantage of the fact that when a building obstructs the line-of-sight between a satellite and a GPS receiver, it causes a detectable drop in the signal to noise ratio (SNR). To demonstrate initial feasibility of our technique, we present the following evidence. Figure 1 demonstrates the SNR values for a satellite as the researcher walks along the side of the Bank of America (BOA) building, occluding the direct line-of-sight path to the satellites with the building. We observe that the SNR drops dramatically when the vector from the GPS receiver to the satellite intersects the target building. These results demonstrate that our assumption—occlusions between the GPS receiver and satellite will lower the SNR—is valid. Our technique combines data samples from many GPS positions to detect buildings and determine their location, size, and height.

2. Data Collection and Feasibility

We initially collected data by walking at a slower than normal pace around the Bank of America (BOA) Plaza in Atlanta, GA, a 55 story skyscraper that is 1023 ft (311.8m) tall. We used two different off-the-shelf consumer grade GPS receivers: a Garmin GPSmap-60CSx and a Garmin GPS 35-LVS. The devices were carried mounted on a wooden support above a backpack. By walking a complete circuit around the building on three days we collected NMEA data samples at a rate of one sample per second which we cropped to use exactly 4000 samples.

After developing our technique, we also collected data from around a different Atlanta skyscraper, the One Atlantic

Center (OAC) building, and verified that the same technique worked with a second set of data gathered at a different location. We gathered data on two days using the Garmin GPS 35-LVS (walking two circuits around the building on each day) to reach 4000 samples. Unlike the slow pace of data sampling for the Bank of America building, when circling the One Atlantic Center building the researcher walked at a medium to fast walking pace, and each circuit took just over fifteen minutes.

Note that for a GPS receiver to achieve a positional fix, it must have a reasonable signal from at least four satellites. In many conditions, GPS receivers detect signals from more than four satellites but do not use the data from satellites with very weak or reflected signals. Our technique makes use of these signals as an indication that the direct line-of-sight path to that satellite may be occluded. Some signals (to satellites that are on or near the horizon) have an elevation that is too low to be of use. We found that ignoring all vectors with an elevation below 15 degrees improved our results significantly. Our techniques make use of statistical averages, so minor errors in GPS position do not greatly harm our results. Our GPS receivers did occasionally drift from the ground-truth path of travel (see traces in Figure 4).

3 Detecting Multiple Buildings

In this section, we demonstrate a method to calculate a *density map* of each site and decide the location of dominant buildings within the testing site. Each NMEA sample collected has multiple signals (each representing a vector to a specific satellite’s elevation and azimuth from the current position of the GPS receiver) that have different SNR values. We convert the vector data into a 2D density map by projecting each vector onto the ground plane. Each vector that has a weak SNR value (possibly due to an occlusion) contributes to the values along a line on the ground plane. Multiple “occluded” vectors passing through the same volume have lines that pass through the same space on the 2D ground plane, and, similar to a Hough transform, probability values combine at the true locations of occlusions while noise values are randomly distributed.

The proceeding overview of our method ignores several key points. Obviously, as we are looking for weak signals (signifying occlusions), we invert the SNR values when adding it to the accumulation buffer. Also, SNRs vary between satellites based upon both their ID and distance. If we add all signal vectors into the buffer directly, the results are biased towards satellites that have overall weaker signals.

To compensate for satellites with different signal levels, we first calculate the maximum, minimum and average SNR values of all signals from each satellite in the data set. We define S_j^{\max} , S_j^{\min} and S_j^{avg} as the maximum, minimum and average SNR respectively of the j th satellite. SNR val-

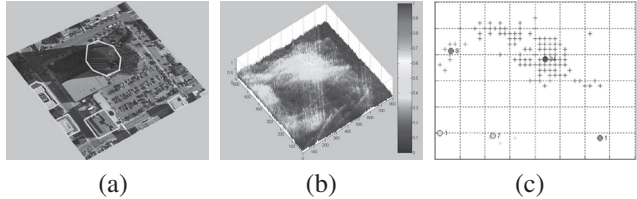


Figure 2. (a) BOA testing site, (b) generated density map, (c) clustering to detect building centers.

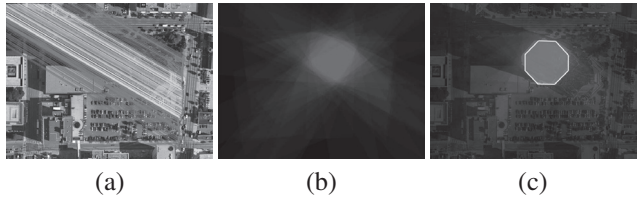


Figure 3. Estimating the region of the BOA building using accumulation of multiple occluded vectors.

ues of individual vectors are labeled as V_{ij} which is the SNR at the i th sample along the traveled path to the j th satellite. Note that all vectors from the i th position share the same real-world origin $i|_{x,y}$, which is determined by the position of the GPS receiver in real-world coordinates. These x, y coordinates, combined with the vectors’ azimuth and elevation, determine the projection of the vector into the accumulation buffer.

At every point in the accumulation buffer, where the vector is projected onto the ground plane, we add the value $L(i, j)$. It is calculated from all samples that fall below the S_j^{avg} as follows:

$$L(i, j) = \begin{cases} \left| \frac{S_j^{\max} - V_{ij}}{S_j^{\max} - S_j^{\min}} \right| & , \text{if } V_{ij} \leq S_j^{avg} \\ 0 & , \text{if } V_{ij} > S_j^{avg} \end{cases}$$

Once we process all signals over every i and j , we normalize the accumulation buffer entries from 0 to 1. Now, the accumulation buffer becomes the *density map*.

Figure 2 shows an overhead view of the Bank of America testing site and the density map generated using our technique. Buildings higher than five stories are marked with green lines in Figure 2(a). The density map shows all five tall buildings with higher than 50% probability. Figure 2(c) shows the clustered result. We applied the mean shift algorithm [3] to find the peak points. These clusters identify candidate regions where buildings are likely to exist. We use the dominant cluster (the blue cluster, which has probabilities larger than 60% on average) as the target building to reconstruct.

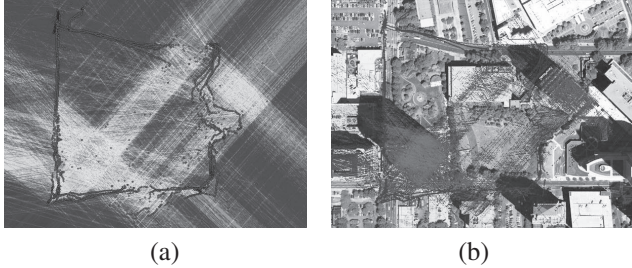


Figure 4. Density map of the OAC site using 4000 samples (red dots) (a), Compared to map imagery (blue dots are sample positions) (b).

4 Building Reconstruction

In this section, we will demonstrate that by using multiple signal vectors passing near and through the dominant cluster target region we can determine the approximate 3D size and shape of the building. As an example of how we estimate the building region, Figure 3(a) shows "occluded" vectors (in green) from a single satellite that pass through a dominant cluster. First we select all vectors that pass through the cluster. Of these vectors, the occluded vectors are those whose SNR falls under a manually selected threshold. By combining all occluded vectors from multiple satellites (Figure 3(b)) we can determine the 2D region of the building. We determine our final building region estimate by selecting the top 10% of the accumulation buffer.

Our result, denoted in red in Figure 3(c), shows our method's estimate of the BOA building's region. The ground truth region is denoted with yellow lines. In this result the difference between the center point of the ground truth region and that of our estimate is only 12 ft (4m). To evaluate our method quantitatively, we calculated the error between our results and ground truth by measuring area difference.

Although visually our result is quite impressive, the actual area difference reports both false-positive regions outside of the ground truth region, and false-negative areas inside of the ground truth region. This result still represents a 23% error in region match. We will analyze the effect that using fewer data samples has on estimation accuracy in Section 5.

Once we estimate the region of a building we can generate its volume based on the collected vectors and their SNRs. We reconstruct a 3D model of the building by using a grid-based voxelization algorithm [8]. As shown in Figure 5, we render voxels that intersect with an occluded vector when they pass above the estimated region.

So far our work has used the law of large numbers to average data samples, removing noise and reducing the ef-

Site	Centroid dist(ft)	Region err(%)	Real height(ft)	Estimated height(ft)	Height err(%)
BOA	7.398	22.73	1023	862.80	15.65
OAC	22.59	22.20	820	705.32	13.98

Table 1. Error results using 4000 samples from each site.

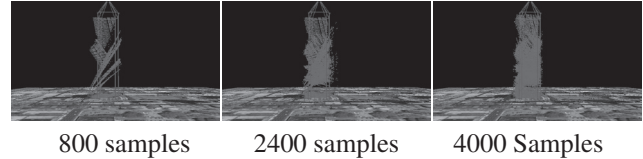


Figure 5. Voxel rendering as the number of data samples increases.

fect of outliers. When using individual vectors to voxelize buildings, near-vertical outliers adversely affect the height estimate. To avoid this problem, we exclude outliers near the top of the building by checking for agreement between at least 10 vectors at each 30 ft increment. This procedure tends to underestimate the height of the building slightly, however, not removing outliers sometimes resulted in a massive height over-estimate. Please view our technical report [7] for more details about our technique.

5 Results and Error Analysis

Our approach produces reasonably accurate results using only 4000 sample points (approximately sixty-six minutes of data collection). Table 1 shows the final result using 4000 samples for both sites. We have determined the building centroids to within 7 and 23 ft when compared to commercially available data sets. Region error approaches 22%, and as shown in Figure 3(c), the visual region is quite close to our comparison data set.

Although the accuracy of building reconstruction may be improved by the use of more sophisticated radio-wave propagation models [10] or better algorithms, it does not appear that our current method can be improved upon by collecting more data samples. The graphs in Figure 6 show the behavior of our error metrics with respect to the number of samples collected. The results clearly show that collecting more data samples markedly improves the results from zero to 1000 samples. An additional 1000 samples (2000 in total) can double the accuracy. Two-thousand samples can be collected in 33 minutes with a consumer GPS device, and in our case, consisted of two circuits around the surveyed area. By the time 2000 samples are used, most results have converged, and additional data produces diminishing returns.

Obviously, the geographical coverage of data samples also plays an important role. In our data collection, 1000 samples corresponds to a complete circuit by the researcher.

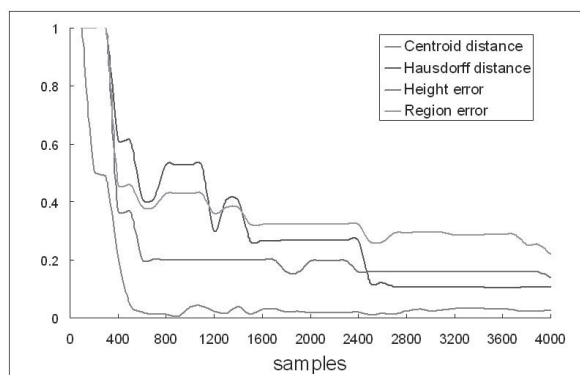


Figure 6. Normalized error metrics as the number of samples increase.

It may be that a smaller number of samples equally spaced along a complete circuit of the area/building under survey will provide equivalent results. Our density map, region estimation, and voxel filling approaches all require widely spaced data around the area/building under examination.

6 Related Work

GPS signals have a long history of being used to detect user motion and make predictions of their goals and context. Additionally, the lack of GPS signals has been used to infer the existence (but not location or size) of tall buildings [6] and as an indoor context detector [2].

Volumetric ray-tracing techniques similar to our approach but using ultrasonics have been used on a small scale with ceiling mounted ultrasonic transmitters and a mobile robot based receiver to build a room-scale environmental model [5]. Volume rendering using data from CT scanners and ultra-sound is widely studied in computer graphics and diverse research efforts already have been proposed [4].

In the only similar work using GPS signals for building detection, Swinford experimented with using non-availability of expected GPS signals to suggest intervening obstacles in the environment. However, this work used GPS data only to add height data to building outlines derived from cartographic maps. Our approach uses only GPS data to detect and measure buildings and does not rely on pre-existing cartographic data. Additionally, Swinford did not present measurements of the accuracy he obtained [9].

7 Conclusion

We have shown that a standard GPS receiver can detect and localize buildings by measuring reduction of SNR caused when an object comes between the receiver and one or more satellites. We introduced an approach that can es-

timate the region of a target building and reconstruct its location, volume and height using only GPS signals. Our approach generated reasonable estimates with around 14-22% errors in region and height. We also demonstrated a method to calculate a density map that predicts where buildings exist in a given area.

This approach is well suited for swarm or crowdsourcing applications, where multiple cooperative agents roam throughout a space sharing GPS data to generate an overview of buildings and other large objects. Even though our approach is less accurate than carefully calibrated computer vision or laser scanning, it has the advantage that GPS receivers are passive, do not require active aiming, automatically self-calibrate, and are inexpensive.

Additional motivation and results including larger figures are available in the associated technical report [7].

Acknowledgements

The authors would like to thank Stéphane Beauregard for his comments on our work and advice.

References

- [1] OpenStreetMap: <http://www.OpenStreetMap.org>.
- [2] D. Ashbrook and T. Starner. Using GPS to learn significant locations and predict movement across multiple users. *Personal Ubiquitous Computing*, 7(5):275–286, 2003.
- [3] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE Pattern Analysis and Machine Intelligence*, 17(5):790–799, 1995.
- [4] T. T. Elvins. A survey of algorithms for volume visualization. *SIGGRAPH 92*, 26(3):194–201, 1992.
- [5] R. K. Harle and A. Hopper. Building world models by ray-tracing within ceiling-mounted positioning systems. In *In Proc' of Ubicomp*, pages 1–17, 2003.
- [6] E. J. Horvitz and J. C. Krumm. Learning, storing, analyzing, and reasoning about the loss of location-identifying signals. US patent 20070005243.
- [7] K. Kim, J. Summet, T. Starner, D. Ashbrook, M. Kapade, and I. Essa. Localization and 3d reconstruction of urban scenes using gps. In *Georgia Institute of Technology Technical Report GT-IC-08-06*, pages 1–8, 2008.
- [8] S. M., B. N., and S. A. Hardware accelerated voxel carving. In *Proc. 1st Iberoamerican Symposium in Computer Graphics*, pages 289–297, 2002.
- [9] R. Swinford. Building on-the-fly world models for pervasive gaming and other ubicomp applications using GPS availability data. In *The IEEE International Workshop on Intelligent Environments*, pages 133–142, 2005.
- [10] J. P. Weiss, P. Axelrad, A. G. Dempster, C. Rizos, and S. Lim. Estimation of simplified reflection coefficients for improved modeling of urban multipath. In *Proc. ION 63rd Annual Meeting*, volume 7, pages 635–643, 2007.